Semiempirical Modeling with Application of Artificial Neural Networks for the Photocatalytic Reaction of Ethylenediaminetetraacetic Acid (EDTA) over Titanium Oxide (TiO_2)

by **Carina A. Emilio**[a]), **Marta I. Litter***[a]), and **Jorge F. Magallanes**[a])[b])

[a]) Unidad de Actividad Química, Centro Atómico Constituyentes, Comisión Nacional de Energía Atómica, Av. Gral. Paz 1499, (1650) San Martín, Prov. de Buenos Aires, Argentina (e-mail: litter@cnea.gov.ar; tel.: 54-11-67727016; fax: 54-11-67727886)

[b]) Departamento de Química Inorgánica, Analítica y Química Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, (1428) Buenos Aires, Argentina

Dedicated to Professor *André M. Braun* on the occasion of his 60th birthday

The use of chemometric techniques, full factorial and *Doehlert* experimental designs, multivariate analysis by MANOVA (= multiple-way analysis of the variance), and artificial neural networks (ANNs) for the photocatalytic reaction of ethylenediaminetetraacetic acid (EDTA) over TiO_2 in aqueous solution is described. Based on the previous knowledge of the photocatalytic system, variables such as EDTA concentration, photocatalyst concentration, pH, and irradiation time were chosen to build a set of experiments for the analysis. By means of MANOVA, the statistical significance of the individual variables and the inspection of interactions between them were analyzed. By the use of ANNs, correlation plots among variables may help to build a semiempirical modeling for understanding and prediction the behavior of the system, optimizing parameters valuable for further technological applications.

Introduction. – Heterogeneous photocatalysis is one of the most studied processes among advanced oxidation technologies. Our laboratory has performed studies on the transformation of different substrates by this technique, working especially on the oxidative degradation of ethylenediaminetetraacetic acid (EDTA) [1–6]. Recently, a complete study of this system has been initiated, establishing the influence of several parameters on the reaction rate such as EDTA concentration, presence of oxidants, *etc.*, nature of the intermediates, and kinetic regime [7][8]. Mechanisms associated with EDTA-photocatalytic degradation are very complicated due to the variety of intermediates formed during the progress of the reaction. In addition, advanced oxidation technologies are characterized by the participation of many experimental variables such as concentration of chemical components, temperature, irradiation wavelength, multiple degradation pathways, among others, which relevantly influence the mechanistic steps. In these complex systems, in which the verification of mechanistic theories and the calculation of kinetic parameters must be made on the basis of a rather large number of experiments, chemometric techniques can be very useful. These techniques allow the appropriate design of the experiments and the diminution of their number, facilitate the interpretation of multivariate phenomena and are valuable tools for scaling up [9].

Artifical neural networks (ANNs) have been shown to be very useful for many chemometric applications, including modeling and optimization. Several advantages

underline the choice of ANNs over other methods: ANNs have the ability to trace the behavior of the variables without the necessity of a hypothesis about the model function. Additionally, ANNs can be successfully applied to nonlinear phenomena [10]. These multivariate analyses are valuable tools that may lead to semiempirical modeling of systems [11]. Chemometrically speaking, semiempirical modeling means that a chemical system is not described by an established mathematical function, but correlation plots among variables allow one to understand the system. This leads to the optimization of the experimental conditions for the design, the scale up of photoreactors, and other parameters useful for further technological applications [10][12][13]. The use of the combination of experimental design (specifically *Doehlert* [14]) and ANNs as analytical and optimization tools was previously proposed [15][16]. Studies for modeling photochemical reactions [9][17], especially photocatalytic [9][18][19] and photo-*Fenton* [16][20–22] systems have been published in recent years.

The present paper describes the use of chemometric techniques, *Doehlert* experimental design, multiple-way analysis of the variance (MANOVA), and artificial neural networks (ANNs) for the photocatalytic reaction of EDTA over TiO_2.

Results and Discussion. – *Choice of Variables and of Experimental Conditions.* To design the type and amount of experiments to be performed for the chemometric analyses, some very well-known features of general heterogeneous photocatalysis and the previous particular knowledge of the EDTA/TiO_2 system were used [7][8][23–26]. Variables with the highest influence on the photocatalytic rate are usually the initial substrate concentration (C_0), the initial pH (pH_0), the amount of photocatalyst, and the irradiation time. Our previous experiments demonstrated that, as usual in heterogeneous photocatalysis, the kinetics of EDTA degradation follows a *Langmuir-Hinshelwood* behavior [7][8] with a saturation of the rate at concentrations higher than 3 mM EDTA. This suggests the influence of the substrate adsorption onto the TiO_2 surface, although this phenomenological behavior could be equally explained by other mechanisms taking place in solution or at the interface. Thus, C_0 was ranged between 1 and 5 mM to cover different regions of the *Langmuir* behavior.

The initial pH of the aqueous solution can significantly affect TiO_2 photocatalysis, and this can be additionally complicated if pH is not controlled during the irradiation. Variations with pH in the driving force for EDTA oxidation over TiO_2 have been claimed to be small [27]. Therefore, changes in the photocatalytic rate can be attributed mainly to changes in the adsorption of the substrate because of the existence of different protolytic EDTA forms (including surface complexes) and changes in the photocatalyst charge in relation with its point of zero charge (pzc) [28]. As it has been suggested that the adsorption density of EDTA over TiO_2 (*Degussa P-25*) did not vary too much in the range 3–5 [23][24], pH_0 was varied between 3.5 and 5.5 in the present case.

Concerning reaction times, only measurements between 30 and 120 min were taken as relevant. The lower limit was chosen to discard large errors frequently found in heterogeneous-photocatalysis systems at shorter times. The upper limit was chosen because previous studies under similar conditions indicated an important depletion of EDTA at that time [7][8].

The experiments were performed with TiO_2 concentrations between 0.5 and 2 g l^{-1}, taking into account that lower values can be detrimental due to an incomplete absorption of light, whereas higher concentrations can cause a 'screening' effect due to high scattering and inappropriate penetration of light [29].

In this preliminary work, other variables such as wavelength and intensity of irradiation, temperature, or amount of oxidants, which can also modify the photocatalytic rate, were kept constant. For example, saturation of the suspension with O_2 or air is crucial for EDTA degradation, as the presence of O_2 prevents the detrimental recombination of photogenerated electrons and holes [7]. Therefore, all experiments were performed under constant O_2 bubbling.

In *Fig. 1*, some selected degradation-averaged profiles are depicted, chosen from a set of experiments used later for the multivariate analyses. The results are representative of how the levels (different values taken by the factors, see p. 647 in [31]) of C_0, TiO_2 concentration, and pH affect the kinetics of EDTA degradation.

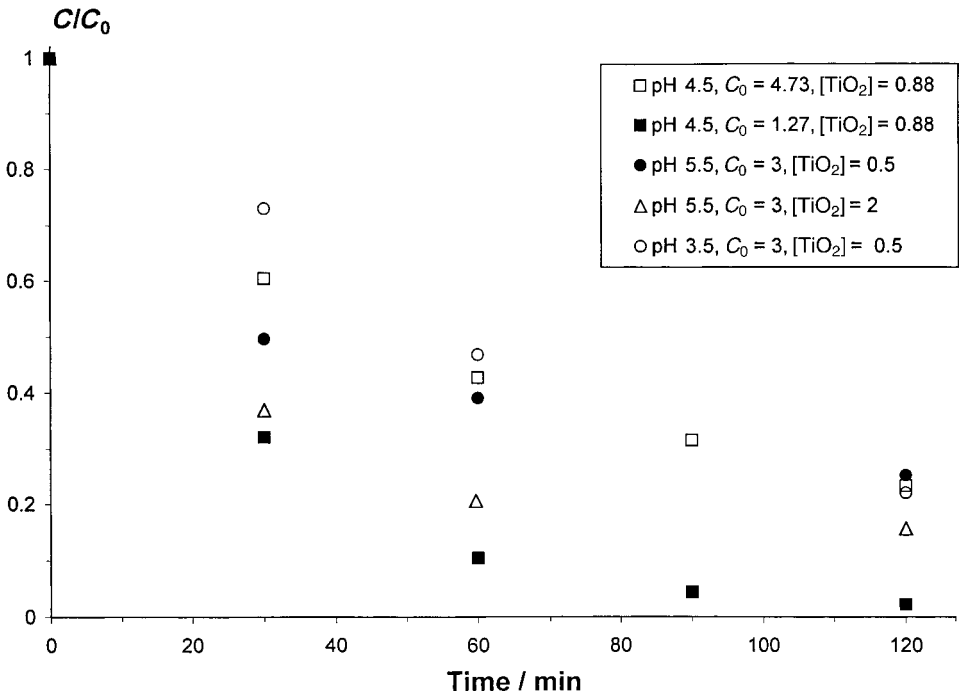

Fig. 1. *Selected profiles of EDTA degradation by photocatalytic oxidation over TiO_2.* Conditions: near UV irradiation (300 nm $< \lambda <$ 500 nm; maximum transmission at 360 nm), $I_0 = 1.1 \cdot 10^{-5}$ Einstein l^{-1} s^{-1}, T 25°. C_0 in mM, [TiO_2] in g l^{-1}.

Experimental errors were estimated around 9, 7, 4, and 2% for 30, 60, 90, and 120 min of irradiation, respectively, based on EDTA depletion, the experimental response. Larger data dispersion was observed for TOC (total organic carbon) results.

Multiple-Way Analysis of the Variance (MANOVA). The first analysis was performed by a three-way MANOVA (see *Appendix*). The data were collected

Table 1. *Full Factorial Experimental Design for MANOVA*

Factors	Levels		
	high	medium	low
$[TiO_2]$/g l^{-1}	2	–	0.5
C_0/mM	1	3	5
pH_0	3.5	–	5.5

Table 2. *Selected Experimental Responses for MANOVA*[a])

Exper.	C_0/mM	$[TiO_2]$/g l^{-1}	pH_0	Degradation extent at 120 min[b])
1	1	0.5	3.5	0.88
2	1	0.5	3.5	0.89
3	3	0.5	3.5	0.78
4	3	0.5	3.5	0.76
5	5	0.5	3.5	0.72
6	5	0.5	3.5	0.72
7	3	2	3.5	0.83
8	3	2	3.5	0.86
9	5	2	3.5	0.83
10	5	2	3.5	0.82
11	1	2	3.5	0.86
12	1	2	3.5	0.84
13	3	0.5	5.5	0.75
14	3	0.5	5.5	0.83
15	5	0.5	5.5	0.74
16	5	0.5	5.5	0.66
17	1	0.5	5.5	0.93
18	1	0.5	5.5	0.86
19	3	2	5.5	0.84
20	3	2	5.5	0.85
21	5	2	5.5	0.76
22	5	2	5.5	0.73
23	1	2	5.5	0.97
24	1	2	5.5	0.98

[a]) Experimental conditions: near UV irradiation ($300\ nm < \lambda < 500\ nm$; maximum transmission at 360 nm), $I_0 = 1.1 \cdot 10^{-5}$ Einstein l^{-1} s^{-1}, T 25°. [b]) EDTA degradation extent defined as $(C_0 - C)/C_0$.

through a full-factorial design [30][31] with 2 levels for pH_0 and $[TiO_2]$ and 3 levels for C_0 (*Table 1*). Duplicate experiments under 12 different conditions were carried out, taking samples after 30, 60, and 120 min of irradiation. Conditions and results at 120 min are shown in *Table 2*.

First inspections of responses verified a normal frequency–behavior distribution of the data population for the 24 experiments with respect to EDTA degradation (*Fig. 2*) and TOC (not shown). This verification of data is necessary to fulfill the theoretical assumptions inherent in MANOVA. However, large data dispersion was observed for TOC results and also for EDTA degradation at 30 min. Therefore, these data were discarded for MANOVA, and only data of EDTA degradation at 60 and 120 min and TOC decrease at 60 min were taken to be reliable. *Table 3* is a typical MANOVA table showing the results of the analysis at 120 min. Taking into account the common level of significance $p < 0.05$, it can be concluded that C_0 and $[TiO_2]$ have a strong influence on

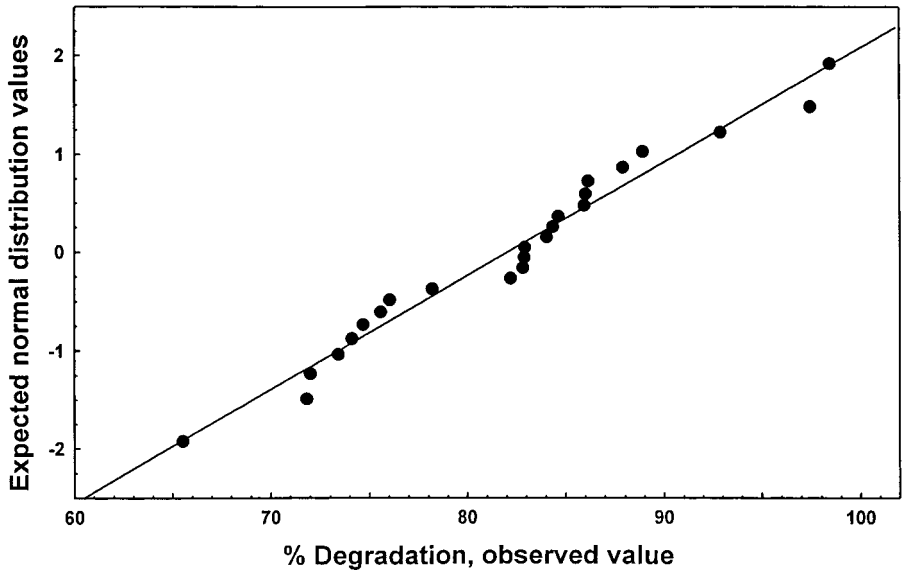

Fig. 2. *Normal probability plot for EDTA degradation data at 120 min by MANOVA*

Table 3. *Results of MANOVA for the Photocatalytic Degradation of EDTA after 120-Min Irradiation* (summary of all effects)[a])

Effect	*df* Effect[b])	*MS* Effect[b])	*df* Error[b])	*MS* Error[b])	*F*[b])	*p*-Level
1	1	4.7838	12	8.954483	0.53423	0.478862
2	1*	188.1005	12	8.954483	21.00629*	0.000629*
3	2*	484.292	12	8.954483	54.08375*	0.000001*
12	1	3.0225	12	8.954483	0.33754	0.572006
13	2*	72.5442	12	8.954483	8.10144*	0.005934*
23	2	14.3448	12	8.954483	1.60196	0.241742
123	2*	43.2769	12	8.954483	4.83299*	0.028868*

[a]) Variables: 1, pH_0; 2, [TiO_2]; 3, C_0. Asterisks refer to statistically significant effects at a level of significance $p < 0.05$. [b]) *df* = degree of freedom; *MS* = mean square; $F = s_1^2/s_2^2$, s_1 and s_2 being the compared standard deviations (see *Appendix*).

the response, but pH_0 exerts a statistically insignificant effect. Furthermore, the results of *Table 3* show an interaction (see *Appendix*) between C_0 and pH_0, which is displayed in *Fig. 3*. The influence of pH_0, which appears to be insignificant in this analysis, is discussed further in the next section. These results were in agreement with those of TOC (not shown). Similar analysis for EDTA degradation after 60 min showed an interaction between C_0 and [TiO_2] but, because of a greater dispersion of data, this result will be examined later.

The interaction among the three variables, shown in *Table 3* and also emerging from TOC results, was not taken into account because it does not have a very significant p-level, reinforced by the insignificant effect of pH.

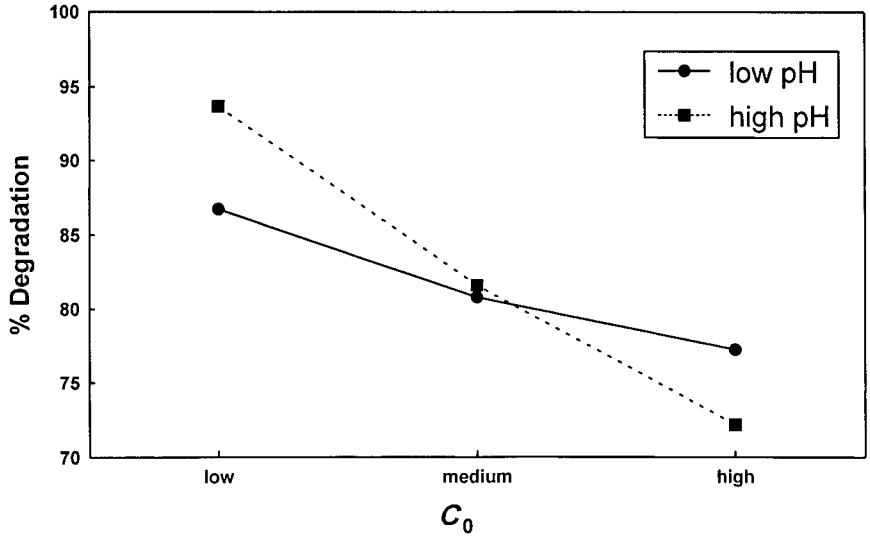

Fig. 3. *Plot of means from two-way MANOVA C_0/pH interaction for the degradation of EDTA*. Experimental data of *Table 2*.

Another four-way MANOVA was performed including time as a new independent variable with levels of 60 and 120 min. The results showed a significant interaction between C_0 and time.

ANNs Modeling. The MANOVA of the system at two levels for pH_0 and $[TiO_2]$ allows to get only linear relationships for the variables. A more detailed description requires more levels and other analytical tools. The interpretation of MANOVA results can be confirmed and enriched by treatment with ANNs, allowing, in addition, the modeling of the system.

The experimental-data collection for the training stage of the network was performed with a *Doehlert* design [14]. In this case, only 13 experiments and their duplicates were enough to cover the whole experimental range (*Table 4*), with 7 levels for C_0, 5 levels for $[TiO_2]$, and 3 for pH_0. In this *Doehlert* design, the highest number of levels is assigned to the variables that need more-detailed descriptions. ANNs Calculations were performed with a home-made program previously reported [11]. The type of net was back propagation of errors (BPE) (see *Appendix* and ref. cit. therein). The most efficient architecture was obtained with a network of 3 layers: 3 input neurons, one hidden layer of 4 neurons, and 4 neurons in the output layer (*Fig. 4*). An additional bias neuron was added to the hidden layer. The optimized response of the network was obtained with a moment $\mu = 0.9$, a learning rate $\eta = 0.5$, and $\alpha = a = 1$ for the transfer function (see *Appendix*, *Eqns. 2* and *3*). The optimal training was attained within approximately 50 epochs; this very fast convergence was due to the efficacy of the experimental design. When the convergence was achieved, the error of prediction kept on a flat minimum up to around 200 epochs, facilitating the selection of the optimum numbers of epochs (see *Appendix*). The ability to make predictions of the network was checked with the set of experiments previously used for MANOVA (*Table 2*),

Table 4. Doehlert *Experimental Design and Responses for ANN Training* [a])

Exper.	C_0/mM	[TiO_2]/g l^{-1}	pH_0	Degradation extent at 120 min	
				1st series	2nd series
1	3.58	1.63	5.3	0.86	0.85
2	1.27	1.63	4.5	0.96	0.96
3	3.00	1.25	4.5	0.95	0.96
4	1.85	1.25	5.3	1.00	1.00
5	3.00	0.50	4.5	0.80	0.82
6	2.42	1.63	3.7	0.97	0.97
7	1.27	0.88	4.5	0.96	0.98
8	4.73	0.88	4.5	0.77	0.77
9	3.00	2.00	4.5	0.95	0.97
10	4.73	1.63	4.5	0.75	0.89
11	2.42	0.88	3.7	0.97	0.98
12	3.58	0.88	5.3	0.85	0.85
13	4.15	1.25	3.7	0.90	0.90

[a]) Conditions: near UV irradiation ($300\,nm < \lambda < 500\,nm$; maximum transmission at 360 nm), $I_0 = 1.1 \cdot 10^{-5}$ Einstein l^{-1} s^{-1}, T 25°.

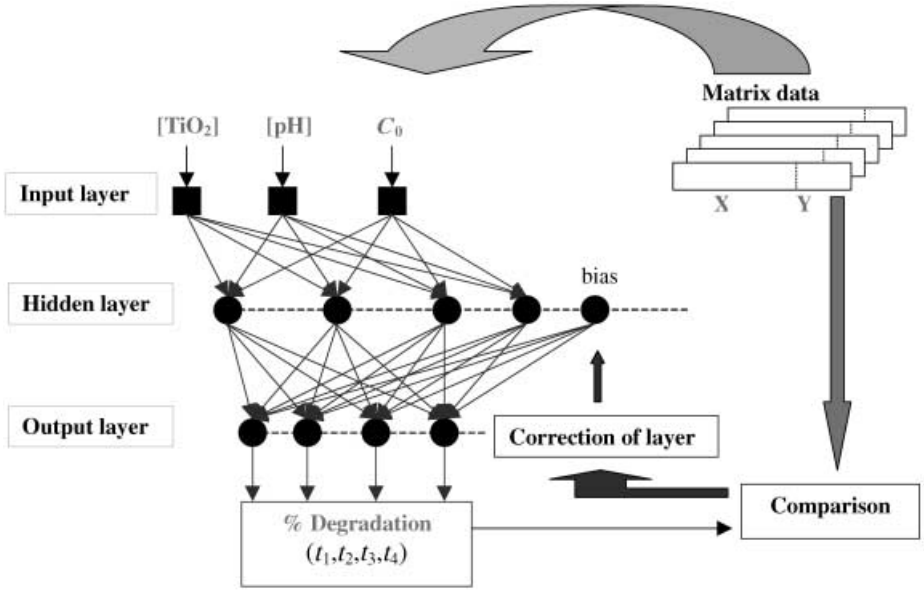

Fig. 4. *Net architecture and training method*

estimating the root mean-square errors (*RMSE*s, see *Appendix*). The *RMSE* was 6% for the training and 7% for the prediction stage, respectively. Note that, as said before, experimental results were affected by a 2–9% error, depending on the irradiation time.

All simulated predictions were performed within the experimental range of the training set (*Table 4*). This means that no extrapolated results were obtained as, so far,

the validity of extrapolated predictions remains uncertain for the ANNs methodology [10][17]. *Fig. 5,a*, shows the agreement of the model for degradation data at 30 min, which have the maximum errors for experiments and predictions. Experimental *vs.* predicted results were compared for all different combinations of the operating conditions. Each experimental point is a combination of variable levels. Considering that the neural network tends to average the dispersion of experimental results, it can be noted that points are distributed around both sides of the line, indicating the goodness of the model. *Fig. 5,b*, displays the behavior of the model for the complete time range of one selected experiment, showing excellent agreement between calculated and experimental points.

Selected response plots corresponding to different analyses obtained by ANNs simulations under particular conditions are presented in *Figs. 6–9*; in the case of *Figs. 7* and *9*, only results at 120 min are displayed, but surfaces of similar shapes were obtained for other reaction times.

As shown in *Fig. 6,a*, at the highest C_0 value, the degradation follows an almost zero-order behavior (with a slight deceleration), but the kinetic regime changes progressively, attaining an order close to one at the lowest concentration. This denotes typical *Langmuir* behavior. However, the observed deceleration seen in all curves could also be due to changes in pH during the reaction and/or to the deactivation of the

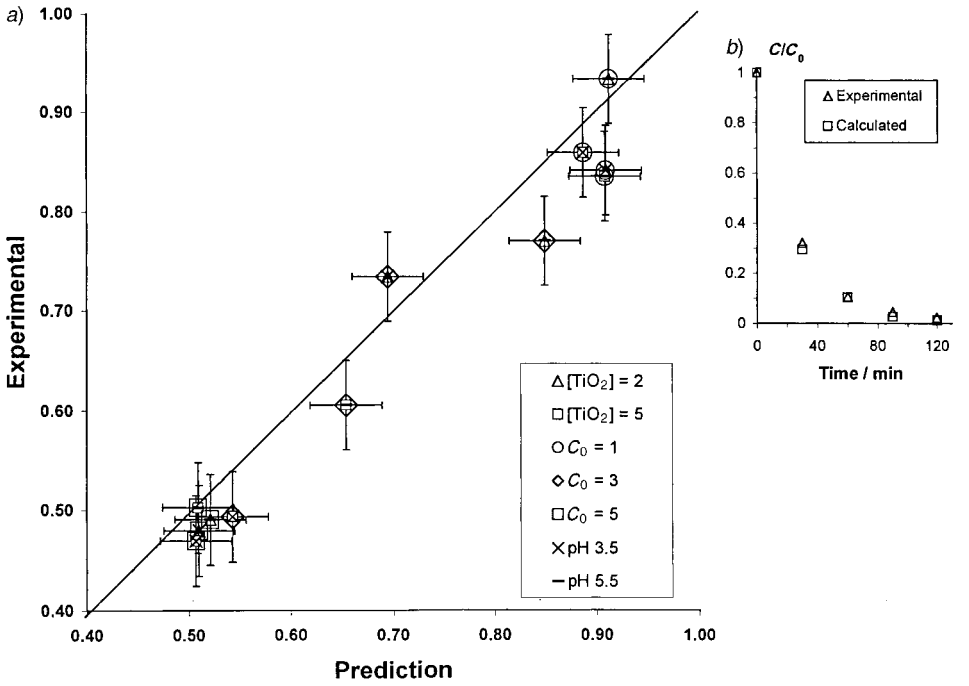

Fig. 5. a) *Correlation between experimental and predicted values simulated by ANNs analysis at 30 min* (error bars taken as 9% for experimental data and 7% for prediction). b) *Experimental and calculated points for* Exper. 7 *of* Table 4. C_0 in mM, $[TiO_2]$ in g l^{-1}.

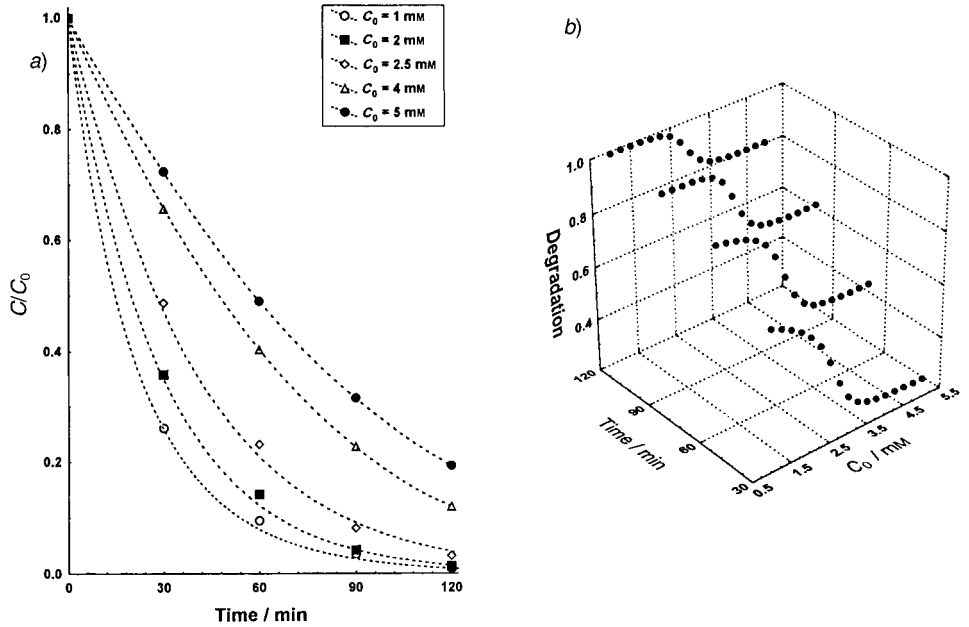

Fig. 6. *Dependence of EDTA degradation as a function of time and* C_0 *predicted by ANNs.* Conditions: $[TiO_2] = 1\ g\ l^{-1}$, pH_0 4.5.

catalyst by products [6]. The change in the kinetic regime is clearly observed in all curves of *Fig. 6, b*, at a $C_0 > 3$ mM, in agreement with our previous results obtained by a classical *Langmuir-Hinshelwood* treatment [7][8]. The interaction between C_0 and time, predicted by MANOVA, appears clearly in the sigmoidal profiles, the height of the step decreasing from shorter to longer times. This reinforces the hypothesis of a *Langmuir* behavior: at $C_0 > 3$ mM, at shorter times, the system stays in the zero-order regime, whereas as long as the reaction progresses, the kinetic regime changes to first order because the substrate is depleted.

In response surface of *Fig. 7*, the interaction between C_0 and pH_0, predicted also by MANOVA, can be observed directly. The profiles of degradation *vs.* C_0 are sigmoidal, with a higher initial slope at the highest pH_0. The interaction can be interpreted by changes in the adsorption density due to the pH-dependent forms of free and complexed EDTA species. This interaction cannot be seen obviously at both limits of the C_0 range, due to total EDTA consumption at 120 min for low C_0, and because of the saturation of the rate at high C_0 due to the increased adsorption density of EDTA onto the TiO_2 surface. The insignificant average effect of pH_0 on the system, which appeared doubtful for MANOVA (see *Table 3*) appears here rather significant at intermediate C_0 values, this result being more reliable because of the more detailed analysis performed by ANNs.

Figs. 8 and *9* show the significant influence of TiO_2 concentration on the response results, as already observed by MANOVA. The increase in conversion with $[TiO_2]$ is consistent with the enhancement of EDTA adsorption in more loaded systems, but this

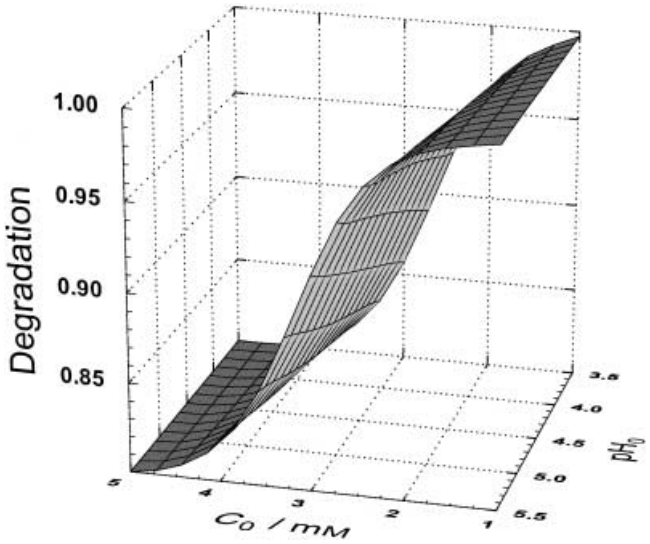

Fig. 7. *Dependence of EDTA degradation as a function of* C_0 *and* pH_0 *at 120 min irradiation time predicted by ANNs.* $[TiO_2] = 1\ g\ l^{-1}$.

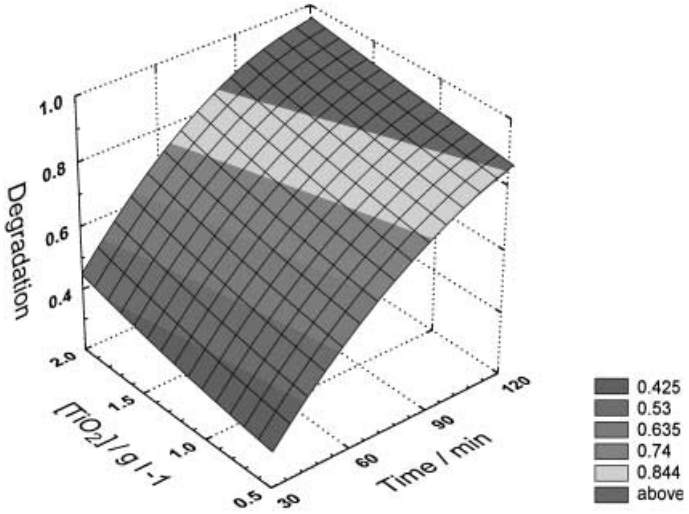

Fig. 8. *Dependence of EDTA degradation as function of* $[TiO_2]$ *and time predicted by ANNs.* Conditions: $C_0 =$ 3 mM, pH_0 4.5.

effect may be attributed also to the lack of a catalyst mass enough to warrant total absorption of light, beyond which the addition of catalyst would not cause any change. Experiments with $[TiO_2]$ higher than $2\ g\ l^{-1}$ (outside of the range investigated here) are needed to prove this hypothesis. *Fig. 8* shows a low but not negligible interaction between $[TiO_2]$ and time, slightly more remarkable in the first stages, whereas *Fig. 9*

displays a weak interaction between [TiO_2] and pH_0 not previously found by MANOVA. Both small interactions can be interpreted by protolytic changes either on the TiO_2 surface or in EDTA speciation, which operate oppositely. As the pH increases, the oxide becomes progressively less positive while EDTA species are more negative; then, the beneficial effect of an increase in pH (below the pzc) can be explained assuming a more important effect of the protolytic equilibrium of the carboxylic acid compared with that of changes in TiO_2. Additionally, *Fig. 9* shows that the behavior of the system changes progressively between the two pH limits; the beginning of a saturation appears in the upper part of the surface from pH_0 *ca.* 4.5 on (a similar effect of saturation has been predicted in [18]). At a fixed [TiO_2], a 10% more of degradation is observed at pH_0 5.5 than at pH_0 3.5; this difference is significant considering that, at 120 min, experimental and prediction errors are very low.

The interaction between pH_0 and time, predicted to be significant by MANOVA, appears doubtful after ANNs analysis (not shown), and can be attributed to the fact that the pH was not controlled. Photocatalytic experiments at constant pH are needed to arrive at a correct conclusion about the role of the pH. Similarly, the apparent interaction between C_0 and [TiO_2], previously found by MANOVA for responses at 60 min, was discarded by the neuronal-network analysis.

It must be remarked that the input EDTA concentration for all multivariate analyses was the initial concentration before TiO_2 addition. The value at the adsorption equilibrium markedly varied with pH, TiO_2 amount, and initial substrate concentration, and an average value was judged to be incorrect to feed the calculations. Studies on EDTA adsorption under all system conditions should be done to interpret its effects.

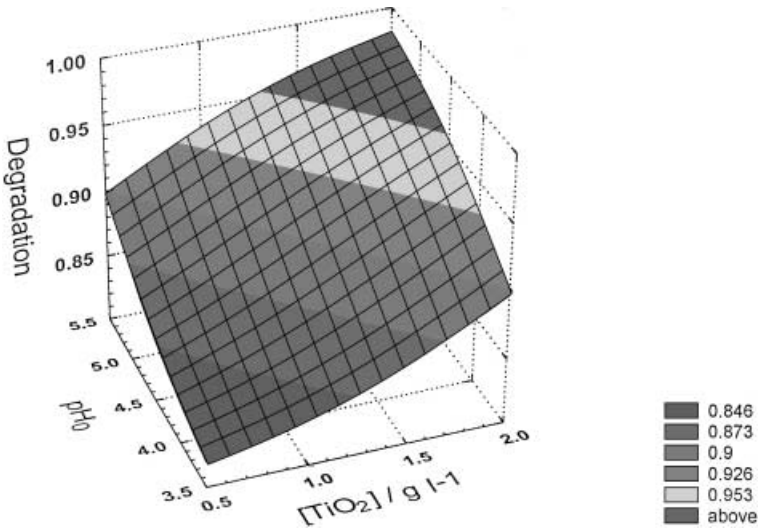

Fig. 9. *Dependence of EDTA degradation as a function of [TiO_2] and pH_0 at 120 min irradiation time predicted by ANNs.* $C_0 = 3$ mM.

Conclusions. – Complex experimental systems difficult to model mathematically and physicochemically can be conveniently handled by means of multivariate analysis. Experimental design and semiempirical modeling offer a support to understand the behavior of the system, and help to develop technological applications. Experimental design permits a lot of experiments and time to be saved and to get a correct registration of the multivariate space covered for the variables. Multilevel experimental designs are necessary to describe the system through response surfaces algorithms greater than second order. The *Doehlert* design here employed gives the highest number of levels to those variables that need more detailed descriptions. Furthermore, it was found that this experimental design helps to achieve a very rapid convergence of the ANN. In addition, it was again demonstrated that BPE-ANNs built by three layers with sigmoidal transfer-function neurons can solve efficiently a chemical system with a remarkably low number of epochs.

The resultant analysis allowed interpretation of the correlations and interactions among variables and verification of previously observed features of the EDTA/TiO_2 photocatalytic system. The main brief conclusions from the studied ranges of C_0, pH_0, and [TiO_2] are *1*) *Langmuir-Hinshelwood* kinetic behavior is confirmed. *2*) The amount of TiO_2 is statistically significant and a higher load is necessary to assure the complete light absorption. There is no interaction between C_0 and TiO_2. *3*) An interaction between C_0 and pH_0 is observed. *4*) In the range 3.5–5.5, the lowest acidities increase moderately the degradation. *5*) A weak interaction of pH_0/TiO_2 exists.

Work is underway to complete this ANNs study, taking into account especially the effects of the pH and the amount of photocatalyst. It is proposed that the described analytical procedure using a selected set of experiments to build a semiempirical model can be a general tool to describe any similar TiO_2 photocatalytic system and to predict its behavior. These multivariate analyses can be a very valuable tool for the optimization of the experimental conditions for technological applications such as design and scale up of photocatalytic reactors, taking into account the role of the influence of the different variables of the system.

Experimental Part

Chemicals. TiO_2 (*Degussa P-25*) was a commercial sample, kindly supplied by the manufacturer (*Degussa AG*, Germany). Na_2EDTA (*Carlo Erba*) was of quality grade and used as provided. All other reagents were at least of reagent grade and used without further purification. Water was doubly distilled in a quartz apparatus. Dil. $HClO_4$ or NaOH solns. were used for initial pH adjustments.

Photocatalytic Experiments. Irradiations were performed in a thermostatted cylindrical *Pyrex* cell irradiated from above by a high-pressure Xe arc lamp (*Osram XBO*, *150*) provided with a bandpass filter (*Schott BG 1*; thickness 3 mm; 300 nm $< \lambda <$ 500 nm; maximum transmission at 360 nm) and a *Schott KG 5* to remove the IR fraction of the incident light. Actinometric measurements were performed by the ferrioxalate method [32]. A photon flow per unit volume of $1.1 \cdot 10^{-5}$ Einstein l^{-1} s^{-1} was calculated.

Photocatalytic runs were done at 25° with, in all cases, 10 ml of a fresh EDTA soln. of known concentration, adjusted to the desired pH_0, and the corresponding mass of TiO_2. The suspension was ultrasonicated for 2 min. Prior to irradiation, suspensions were stirred in the dark for 30 min to assure adsorption equilibration. Irradiations were performed under magnetic stirring, and a H_2O-sat. O_2 stream was bubbled in the suspension at a 0.2 l min^{-1} constant rate throughout the experiment. During the irradiation, the pH was not controlled, and it increased along the experiments due to the formation of basic intermediates including ammonia [7][8]. These pH changes depended on the starting value, the difference being higher (2–3 units) under the most acidic conditions. Samples were withdrawn periodically and filtered through 0.22-μm *Millipore* filters. The EDTA

concentration was evaluated by spectrophotometric analysis with bis(2,4,6-tripyridyl-1,3,5-triazine)iron(II) in the VIS range [33]. UV/VIS-Absorption measurements were performed with a *Shimadzu 210A* spectrophotometer. Total organic carbon (TOC) was measured with a *Shimadzu 5000A* TOC analyzer.

This work was performed as part of the *Comisión Nacional de Energía Atómica CNEA-CAC-UAQ* project #95-Q-03-05, *Agencia Nacional de Promoción de la Ciencia y la Tecnología* (*ANPCYT*) project PICT98-13-03672, *Consejo Nacional de Investigaciones Científicas y Técnicas* (*CONICET*) project PIP662/98, and *Secretaría para la Tecnología, la Ciencia y la Innovación Productiva* project ES/PA/00-EXIII/005. We thank the *CYTED* Program, a 'Scientific Iberoamerican Cooperation Action' (*CYTED VIII-G*), for subsidy. *C. A. E.* thanks *ANPCYT* for a fellowship to perform this work. *M. I. L.* is a member of *CONICET*.

Appendix. *Basic Concepts about MANOVA.* The statistical significance of a result is an estimated measure of the degree to which it is true or representative of the population of data. The variability of a set of data is partially due to random errors, but if a factor introduces some influence in the measurements, this influence will appear like a systematic error. Then, the analysis of the variance (ANOVA) helps to find out the possible influence of a factor on the measurements. According to *Scheffe* [34], ANOVA is defined as a 'statistical technique for analyzing measurements that depends on several kinds of effects operating simultaneously to decide which kind of effects are important and to estimate the effects'. When only a single possible effect is studied, the procedure is called one-way ANOVA, but if several factors are being analyzed at the same time, the process is called multiple-way ANOVA or MANOVA.

To analyze whether a factor has an effect on the measurements (the responses), a statistical hypothesis test is usually applied, whose description is beyond the scope of this work and can be found in basic texts of statistics. Usually, a table is built containing the calculations and parameters used to interpret the ANOVA or MANOVA results. The significance of these parameters can also be found in statistics texts. In our case, *Table 3* shows the results of MANOVA analysis, displaying the following parameters: the degrees of freedom (*df*), the mean-square (*MS*) of errors and effects, the *F*-test ($F = s_1^2/s_2^2$, where s_1 and s_2 are the compared standard deviations, with $s_1 > s_2$), and the significance level (*p*-level). Specifically, the *p*-level represents the probability of error involved in accepting the validity of an observed value as representative of the population. In many areas of research, a *p*-level < 0.05 is an acceptable probability level.

When the effect of one factor is modified by the level of another, it is said that there is an 'interaction' between these factors. Interactions can occur among more than two factors.

Theoretical Aspects of the ANNs Technique. The ANNs try to mimic a natural neural network and use a scheme composed of neurons and channels to propagate information. The back propagation of errors (BPE) net architecture can be built by multiple layers of neurons, but the assembling of three layers, one for input, a hidden layer, and an output (*Fig. 4*) is enough to solve a variety of problems. The number of neurons of the input layer is determined by the number of input variables and, analogously, the number of neurons of the output layer is determined by the number of output variables. The complete architecture, *i.e.*, the number of hidden layers with the corresponding number of neurons in each one, is usually determined by trial and error. Connections in the net are established only among different layers.

During the training of the network, the information is collected by the input layer, where it is distributed to all the neurons of the hidden layer. A weighed sum (*Net*) of the input signals arrives to the neuron, and can be described by *Eqn. 1*, where *l* stands for layer, *j* for neuron, and *i* for inputs, *m* being the total number of inputs of the neuron. In the present case, the starting weights *w* for the training stage were randomly established in the range $-1 < w < 1$. The Nets values are then modified by a mathematical transfer function. As this function can be linear or nonlinear, the ANNs have the ability to respond also to nonlinear phenomena. In our case, we used the nonlinear sigmoidal function of *Eqn. 2*, where α and a are constants related to the form of the transfer function [10].

$$Net_j^l = \sum_{i=1}^{m} w_i^l \cdot x_i^l \tag{1}$$

$$Out(Net) = [1 + \exp(\alpha(a - Net))]^{-1} \tag{2}$$

The derivative of the transfer function (*Eqn. 2*), on which calculations depend, has a general expression of the form $f'(x) = f(x) \cdot (1 - f(x))$, whose simplicity allows a very fast calculation (see below).

Subsequently, the information is transferred to the following layer, where a similar process yields new output, which can be the input for another layer or the final result.

In the training stage, the net output is compared with the true (measured) value of a 'learning set' of experiments, and the weights *w* are corrected to improve the output until a convergence is achieved. During the training, the data matrix is passed through the net several times, and each time is called an 'epoch'. The adjustments of the weights w_i are made using a 'gradient descent' algorithm, whose most general form is given by *Eqn. 3*. The parameter η, called learning rate, determines the speed of change of the weights, while the momentum (μ) regulates the most recent correction to prevent sudden changes in the adjust direction. The values of η and μ must be empirically adjusted to optimize the efficiency of the calculation.

$$\Delta w_{ji}^{l} = \eta \cdot \delta_j^l \cdot Out_i^{l-1} + \mu \cdot \Delta w_{ji}^{l\,(\text{previous})} \tag{3}$$

The concept of convergence means that, after several epochs (usually thousands), the net approaches progressively the expected true values within a reasonable error deviation. During the calculations, the convergence cannot be assured, and sometimes it is not attained. The error of the convergence is checked by a so-called 'cost function', which, in our case, is given by *Eqn. 4*, *i.e.* the sum of the quadratic differences between the net outputs and the experimental results *y*, where *n* is the number of vector inputs or 'objects'. This error depends also on η and μ.

$$RMSE = \sqrt{1/n. \sum_{i=1}^{n} (y_i - Out_i)^2} \tag{4}$$

Once the net has been trained, the weights *w* remain unchanged, and their ability for predictions is checked against a new set of experiments ('test set'). Usually, errors during training descend continuously through successive epochs, but errors during predictions, after an initial lowering, begin to increase again. The number of epochs appropriate to stop the training is a compromise between these two kinds of errors. The *RMSE* calculated from data of the 'test set' indicates the ability of prediction of the ANN and its accuracy.

The complete description of the calculation and programming of the BPE network is beyond the scope of this report. A specific discussion of this subject can be found in [10][35–37].

After training the net and having attained an acceptable accuracy, it is possible to simulate any experiments by introducing the corresponding set of values of the input variables to obtain the corresponding responses. In this way, it is possible to study the behavior of the variables and the relationships among them.

REFERENCES

[1] M. I. Litter, J. A. Navío, *J. Photochem. Photobiol. A: Chem.* **1994**, *84*, 183.
[2] J. A. Navío, G. Colón, M. I. Litter, G. N. Bianco, *J. Mol. Calc.* **1996**, *106*, 267.
[3] J. A. Navío, J. J. Testa, P. Djedjeian, J. R. Padrón, D. Rodríguez, M. I. Litter, *Appl. Catal. A: General* **1998**, *178*, 191.
[4] E. A. San Román, J. A. Navío, M. I. Litter, *J. Adv. Oxid. Technol.* **1998**, *3*, 261.
[5] S. Botta, G. M. Restrepo, J. A. Navío, M. I. Litter, *J. Photochem. Photobiol. A: Chem.* **1997**, *129*, 89.
[6] U. Siemon, D. Bahnemann, J. J. Testa, D. Rodríguez, N. Bruno, M. I. Litter, *J. Photochem. Photobiol. A: Chem.* **2002**, in press.
[7] P. A. Babay, C. A. Emilio, R. E. Ferreyra, E. A. Gautier, R. T. Gettar, M. I. Litter, *Water Sci. Technol.* **2001**, *3*, 193.
[8] P. A. Babay, C. A. Emilio, R. E. Ferreyra, E. A. Gautier, R. T. Gettar, M. I. Litter, *Int. J. Photoenergy* **2001**, in press.
[9] A. M. Braun, L. Jacob, E. Oliveros, C. A. O. Do Nascimento, 'Up-Scaling Photochemical Reactions', in 'Advances in Photochemistry', Eds. D. H. Volman, G. S. Hammond, D. C. Neckers, Wiley, New York, 1993, Vol. 18, p. 235.
[10] J. Zupan, J. Gasteiger, 'Neural Networks for Chemists', VCH Publishers, New York-Weinheim, 1993.
[11] J. F. Magallanes, P. Smichowski, J. J. Marrero, *Chem. Inf. Comput. Sci.* **2001**, *41*, 824.
[12] O. Jiménez, I. Benito, L. M. Marina, *Anal. Chim. Acta* **1997**, *353*, 367.
[13] R. Goodacre, É. M. Timmins, A. Jones, D. B. Kell, J. Maddock, M. L. Heginbothom, J. T. Magee, *Anal. Chim. Acta* **1997**, *348*, 511.
[14] D. H. Doehlert, *Appl. Statit.* **1970**, *19*, 231.
[15] E. Oliveros, F. Benoit-Marquié, E. Puech-Costes, M. T. Maurette, C. A. O. Nascimento, *Analusis* **1998**, *26*, 326.

[16] S. Göb, E. Oliveros, S. H. Bossmann, A. M. Braun, R. Guardani, C. A. O. Nascimento, *Chem. Eng. Proc.* **1999**, *38*, 373.
[17] C. A. O. Do Nascimento, E. Oliveros, A. M. Braun, *Chem. Eng. Proc.* **1994**, *33*, 319.
[18] L. Jakob, E. Oliveros, O. Legrini, A. M. Braun, in 'Photocatalytic Purification and Treatment of Water and Air', Eds. D. F. Ollis, H. Al-Ekabi, Elsevier Science Publishers, The Netherlands, 1993, p. 511.
[19] E. Bessa, G. Lippel Sant'Anna Jr., M. Dezotti, *J. Adv. Oxid. Technol.* **1999**, *4*, 196.
[20] S. Göb, E. Oliveros, S. H. Bossmann, A. M. Braun, R. Guardani, C. A. O. Nascimento, *J. Inf. Recording* **2000**, *25*, 447.
[21] E. Oliveros, S. Göb, S. H. Bossmann, A. M. Braun, C. A. O. Nascimento, R. Guardani, 'Proceedings of the Third Asia Pacific Conference', Eds. X. Hu, P. L. Yue, World Scientific, Singapore, 2000, p. 577.
[22] S. Göb, E. Oliveros, S. H. Bossmann, A. M. Braun, C. A. O. Nascimento, R. Guardani, *Water Sci. Technol.*, in press.
[23] D. N. Furlong, D. Wells, W. H. F. Sasse, *J. Phys. Chem.* **1985**, *89*, 626.
[24] D. N. Furlong, D. Wells, W. H. F. Sasse, *J. Phys. Chem.* **1985**, *89*, 1922.
[25] D. N. Furlong, D. Wells, W. H. F. Sasse, *Aust. J. Chem.* **1986**, *39*, 757.
[26] D. N. Furlong, D. Wells, W. H. F. Sasse, *J. Phys. Chem.* **1986**, *90*, 1107.
[27] D. N. Furlong, D. Wells, W. H. F. Sasse, *Aust. J. Chem.* **1986**, *39*, 757.
[28] D. Bahnemann, J. Cunningham, M. A. Fox, E. Pelizzetti, P. Pichat, N. Serpone, in 'Aquatic and Surface Photochemistry', Eds. G. R. Helz, R. G. Zepp, and D. G. Crosby, Lewis Publ., Boca Raton, Florida, 1995, p. 261.
[29] J. Blanco, S. Malato, 'Solar Detoxification', UNESCO Engineering Learning Package 2000 (to be published); fully available at: http://www.unesco.org/science/wsp/publications/solar.htm.
[30] D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte, L. Kaufman, 'Chemometrics: a Textbook', 2nd edn. Elsevier, The Netherlands, 1988.
[31] D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. Smeyers-Verbeke, 'Data Handling in Science and Technology', 'Handbook of Chemometrics and Qualimetrics, Part A', Elsevier, Amsterdam, 1997, Vol. 20A.
[32] C. G. Hatchard, C. A. Parker, *Proc. Roy. Soc. (London) A* **1956**, *235*, 518.
[33] B. Kratochwil, M. C. White, *Anal. Chem.* **1965**, *37*, 111.
[34] H. Scheffe, 'The Analysis of Variance', Wiley, New York, 1959.
[35] Y. Chauvin, D. Rumelhart, 'Backpropagation: Theory, Architectures and Applications', Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey, 1995.
[36] L.-M. Fu, 'Neural Networks in Computer Intelligence', McGraw Hill, Singapore, 1994.
[37] 'Backpropagation: Theory, Architecture, and Applications', Eds. Y. Chauvin and D. E. Rumelhart, Lawrence Erlbaum Associates, Inc., Publish., Hillsdale, New Jersey, 1994.

Received June 15, 2001